

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NPTEL

NPTEL ONLINE CERTIFICATION COURSE

Nonlinear Programming

Lec – 05

Convex Programming Problems

Dr.S.K. Gupta

Department of Mathematics

Indian Institute of Technology Roorkee

Welcome to lecture series on nonlinear programming in the previous lectures we have seen that what convex functions are what are their properties. Now we will see some convex programming problems, what they are, and how are they important that we will see in this lecture. So first what are convex programming problem. Now consider this general problem, general mathematical programming problem that is minimizing the function effects subject to GIX less than or equal to 0 if 1 to m. That means we are having m number of constraints.

(Refer Slide Time: 00:56)

Convex Programming Problem
The optimization problem of the form:

$$\begin{aligned} & \text{Min } f(x) \\ & \text{subject to: } g_i(x) \leq 0, \quad i = 1, 2, \dots, m. \end{aligned}$$

is called a convex programming problem (CPP) if f and g_i , ($i = 1, 2, \dots, m$) are convex functions.

Theorem
Let g_i for each $i = 1, 2, \dots, m$, be a convex function. Then

$$S = \{x \in \mathbb{R}^n, g_i(x) \leq 0, i = 1, 2, \dots, m\}$$

is a convex set.

IT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 2

So this problem is called convex programming problem if the function F which is the objective function and G_i for all i from 1 to m are convex functions. So our problem which is of minimizing type and all constraints less than equal to is called convex programming problem if the objective function F and all the constraints G_i are convex, so it is called such problems are convex programming problems. Now if G_i for each i from m to be convex function then this set is always a convex set.

(Refer Slide Time: 01:43)

$$S = \left\{ x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, 2, \dots, m \right\} \quad g_i \rightarrow \text{convex function (if)}$$

Let $x_1, x_2 \in S \Rightarrow g_i(x_1) \leq 0 \quad \forall i$
 $g_i(x_2) \leq 0 \quad \forall i$

$x = \lambda x_1 + (1-\lambda)x_2, \lambda \in [0,1]$

$g_i(x) = g_i(\lambda x_1 + (1-\lambda)x_2)$
 $\leq \lambda g_i(x_1) + (1-\lambda)g_i(x_2) \quad (\because g_i(x) \text{ are convex function})$
 $\leq \lambda \cdot 0 + (1-\lambda) \cdot 0 = 0$

$\Rightarrow g_i(x) \leq 0 \quad \forall i$
 $\Rightarrow x \in S$

Now let us see that what this set is, the set S is nothing but all X belongs to \mathbb{R}^n such that $G_i(X)$ less than equals to 0, i from 1 to M where G_i all G_i are convex functions, all G_i are convex functions for all i , then if all G_i are convex function then this set consisting of all X belongs to \mathbb{R}^n such that $G_i(X)$ less than equals to 0 for all i , then this set is a convex set, proof is very easy we can simply, we can prove it.

So in order to prove that is a convex set let two points, two arbitrary points of an X_2 less than S and what we have to show we have to prove that the convex linear combination of these two arbitrary point is also in S . Then we can say that this set is the convex set, so these two in S this means $G_i(X_1)$ will be less than equals to 0 for all i , and $G_i(X_2)$ is also less than equals to 0 for

all I . Now take convex linear combination of these two points, so this is the convex linear combination λ between 0 and 1 okay.

Now to show that this set S is a convex set we have to prove that this X is in S , so how we can show this you take $G_I(X)$ what it is, it is nothing but $G_I(\lambda x_1 + (1-\lambda)x_2)$ and since G_I are convex functions for every I so this means that this is less than equals to $\lambda(G_I x_1 + (1-\lambda)(G_I x_2))$ since G_I are convex for all I okay, are convex functions. Now $G_I(x_1)$ is less than or equal to 0 for all I and $G_I(x_2)$ is also less than equal for all I .

So this is less than equals to $\lambda(0) + (1-\lambda)(0)$ which is equals to 0. So this implies that $G_I(x)$ is less than equals to 0 for all I and this implies X belongs to this. So hence S is a convex set okay. So what I want to say that if you take all the constraint less than equal to type and all constraints are convex function, then the collection of all those X which satisfy all the constraint of $G_I x$ less than equal to 0 types is always a convex set okay. Now let us discuss few examples of convex programming problems.

(Refer Slide Time: 04:59)

$\text{Min } x_1 + x_2$ s.t. $x_1^2 + x_2^2 \leq 1 \rightarrow g_1(x) = x_1^2 + x_2^2 - 1 \leq 0$
 $x_1^2 \leq x_2 \rightarrow g_2(x) = x_1^2 - x_2 \leq 0$

$\nabla^2 g_1 = \begin{pmatrix} \frac{\partial^2 g_1}{\partial x_1^2} & \frac{\partial^2 g_1}{\partial x_1 \partial x_2} \\ \frac{\partial^2 g_1}{\partial x_1 \partial x_2} & \frac{\partial^2 g_1}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

$\lambda = 2 > 0 \Rightarrow \det = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 > 0 \Rightarrow g_1$ is convex function

$\nabla^2 g_2 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \quad \lambda_1 = 2, \lambda_2 = 0$

\rightarrow positive semi-definite
 $\Rightarrow g_2$ is convex function.

First example is minimizing x_1+x_2 subject to that the first constraint is $x_1^2+x_2^2$ less than equal to 1 and the second constraint is x_1^2 less than equal to x_2 . Now this constraint can be rewritten as $G_1x=x_1^2+x_2^2-1$ less than equal to 0 and it is can be written as the second constraint $G_2x=x_1^2-x_2^2$ which is less than equal to 0 okay. And of course the objective function is the objective is minimizing $F=x_1+x_2$ okay.

Now we already know that all linear functions are convex, so this function x_1+x_2 is obviously convex okay, it is a linear function and linear function are always convex as well as concave. So this function is convex. Now if you talk about the first constraint to show that whether it is convex or not you take hessian matrix of G_1 , what is the hessian matrix of G_1 it will be nothing but $(\delta^2G_1/\delta x_1^2 \quad \delta^2G_1/\delta x_1\delta x_2 \quad \delta^2G_1/\delta x_2\delta x_1 \quad \delta^2G_1/\delta x_2^2)$ and this is nothing but if you take second derivative of G_1 respect to x_1 it is 2 there is no term involving x_1, x_2 in this constraint.

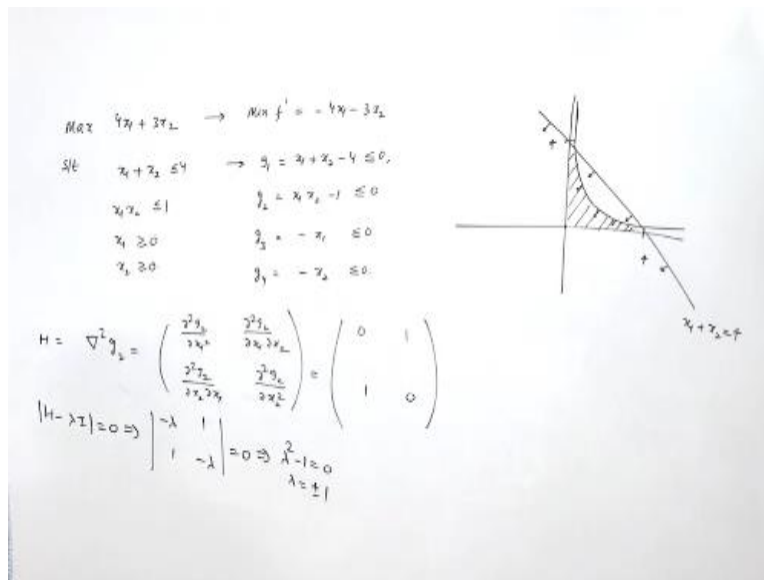
So this is 0, this is 0, this is 2 again 2. So this is the hessian matrix respect to G_1 , now to see whether it is convex function or not we take the leading principal minor okay, we use deck test, so what is D_1 , D_1 is 2 which is greater than 0, of course D_1 is 2 greater than 0, and D_2 is nothing but it is determinant of $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ which is nothing but 4 again greater than 0. Since D_1 and D_2 both are greater than 0 this means hessian matrix is, this means hessian matrix positive definite and this means function is convex okay.

So this means, this implies D_1 is convex function, or you can use eigenvalue test, you see eigenvalues of this matrix are nothing but 2 and 2, both eigenvalues are syndicate and 0 so this means function is strictly convex adherence convex okay. Now similarly you take second constraint hessian matrix second constraint is nothing but it is $\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ even in the same structure we can find the initial matrix of G_2 and now the eigenvalues are, eigenvalues of this is nothing but eigenvalue is nothing but 2 and 0 we can simply see because the diagonal matrix.

And eigenvalues are greater than equal to 0 so this means this matrix is positive semi-definite enhance the function G_2 is convex function okay. So this means this is positive semi-definite and this implies G_2 is convex function. So hence we have seen that for this problem for particular example E1 the function F which is objective function is convex constraint G_1 is convex,

constraint G2 is convex, hence it is a convex programming problem, because the involved objective function all the constraints are convex hence this is a convex programming problem. In the second example objective functions are maximizing type.

(Refer Slide Time: 09:26)



But it is linear you see what will be this objective function maximizing $4x_1 + 3x_2$ subject to what are the constraint it is $x_1 + x_2$ less than equals to 4, it is $x_1 x_2$ less than equals to 1, x_2 greater than equal to 0, and x_1 greater than equal to 0. Now this objective function can be recast as minimum of some f which is nothing but $-4x_1 - 3x_2$ okay, this is same as you can maximizing you can write as minimizing of negative of this function okay.

And subject to whatever conditions, condition will be this will be nothing but G1 which is nothing but $x_1 + x_2 - 4$ less than or equal to 0, the second constraint will be $G_2 = x_1 x_2 - 1$ less than equals to 0, third constraint will be nothing but $-x_1$ which is less than equal to 0, and G4 will be nothing but $-x_2$ which is less than equal to 0. Because for convex programming problem the format of the problem should be of minimizing type and all constraint must be less than equal to 0 okay.

Now this is minimizing function and this is linear hence convex because all linear functions are convex, the first constraint is linear it is convex, obviously the third constraint is linear convex, fourth constraint is linear and hence convex. Now we have to check only for G2 constraint whether it is convex or not, if it is convex that will be a convex programming problem otherwise it will not be a convex programming problem.

Now to check whether g2 is convex or not find out hessian matrix of g2 what will be hessian matrix of g2 the hessian matrix of g2 will be it is $\partial^2 g_2(\partial x_1^2) \partial^2 g_2(\partial x_1, \partial x_2)$ it involves only two variable hence only two cross two matrix would be there, so this is the hessian matrix this is nothing but it is 0, 1 it is 1 and it is 0, so this will be hessian matrix respective to g2 now if you find Eigen value of these matrix.

So how to find Eigen values of this matrix suppose it is suppose it H, so $|H - \lambda I|$ determinant should be 0 and this implies $|\lambda - 1, 1 - \lambda|$ should be 0 this implies $\lambda^2 - 1 = 0$ $\lambda = 1, -1$, so for some λ are positive and some λ are negative one λ is positive one λ is negative, that means the function is not convex because for convexity all λ must be ≤ 0 that means function is convex, hence this constrained is not convex therefore this problem is not convex programming problem.

Because if it is a convex programming problem then objective function as well as all constrained must be convex okay you can see it graphically also see what is a first constrained it is $x_1 + x_2 \leq 4$, this 4 this 4, this is a first constrained $x_1 + x_2 = 4$, and ≤ 4 means shade to it is origin because if you take 0, 0 so $0 + 0 \leq 4$ a true it wholes, so we shade towards origin, okay now if you take the second constrained $x_1, x_2 \leq 1$ it is rectangular hyperbola okay $x_1, x_2 \leq 1$ that means this type of hyperbola we should hiving we might be having.

And now it takes 0, 0 is 0 which is ≤ 1 that is true that shades towards origin again and x_1, x_2 both are no negative that is an that means the first quadrant that we are talking only the first position of the rectangular hyperbola because we need only that position which is a first quadrant okay, so that means this is the region which is below this line and below this hyperbola, so this is the region and we can easily see that this region is not convex set, and from the previous theorem it must be a convex set.

(Refer Slide Time: 14:21)

Convex Programming Problem

The optimization problem of the form:

$$\begin{aligned} & \text{Min } f(x) \\ & \text{subject to: } g_i(x) \leq 0, \quad i = 1, 2, \dots, m. \end{aligned}$$

is called a convex programming problem (CPP) if f and g_i ($i = 1, 2, \dots, m$) are convex functions.

Theorem

Let g_i for each $i = 1, 2, \dots, m$, be a convex function. Then

$$S = \{x \in R^n; g_i(x) \leq 0, i = 1, 2, \dots, m\}$$

is a convex set.

BT BPOURKEE NPTEL ONLINE CERTIFICATION COURSE 2

If all g_i 's are convex function then this set must be a convex set, however this is not a convex set, so we can easily say that since the region is not a convex set so it is sort of convex programming problem okay.

(Refer Slide Time: 14:37)

Different formats of the problem (CPP)	
Optimization problem	Conditions for (CPP)
Min $f(x)$ subject to: $g_i(x) \leq 0, (i = 1, 2, \dots, m)$	f and $g_i (\forall i)$ are convex.
Max $f(x)$ subject to: $g_i(x) \leq 0, (i = 1, 2, \dots, m)$	f is concave and $g_i (\forall i)$ are convex.

Now we have different formats of convex programming problem let us see the first format we have already discussed that is a minimizing type functions subject to all constrain \leq this will be a convex programming problem, if function f and all g_i 's are convex the second type if the objective function is a maximizing type and all constrains are \leq type then it will be a fugues programming problem if function is concave and all constrains are convex.

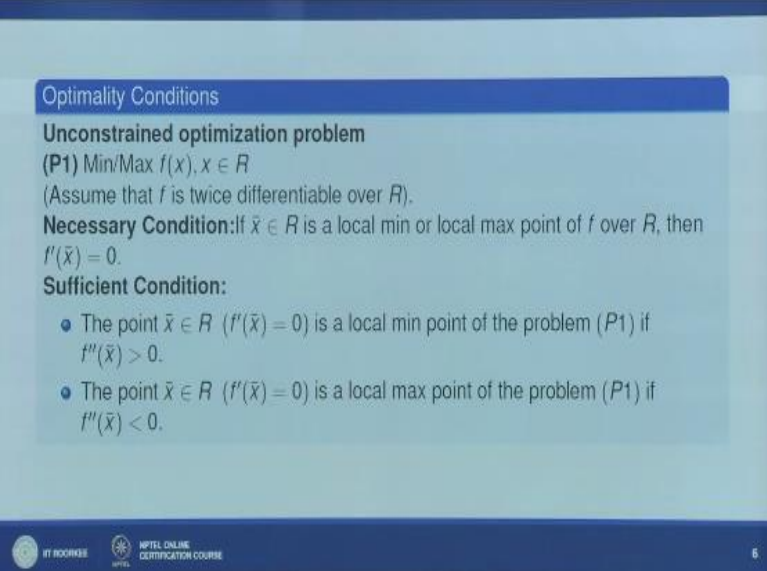
(Refer Slide Time: 15:13)

Min $f(x)$ subject to: $g_i(x) \geq 0, (i = 1, 2, \dots, m)$	f is convex and $g_i (v_i)$ are concave.
Max $f(x)$ subject to: $g_i(x) \geq 0, (i = 1, 2, \dots, m)$	f and $g_i (v_i)$ are concave.

IT BOARDS NPTEL ONLINE CERTIFICATION COURSE 5

The third form is function is of minimizing type objective function and all constrains are \geq to type then it will be a convex programming problem if function is convex and all constrains are concave, okay and if we have maximizing type objective function in all constrains are \geq type then it will be a convex programming problem if and all g_i 's are concave. So these are the 4 different formats of a convex programming problem, so by which we can see whether the problem is convex are not these are the conditions.

(Refer Slide Time: 15:53)



Optimality Conditions

Unconstrained optimization problem
(P1) Min/Max $f(x)$, $x \in R$
(Assume that f is twice differentiable over R).

Necessary Condition: If $\bar{x} \in R$ is a local min or local max point of f over R , then $f'(\bar{x}) = 0$.

Sufficient Condition:

- The point $\bar{x} \in R$ ($f'(\bar{x}) = 0$) is a local min point of the problem (P1) if $f''(\bar{x}) > 0$.
- The point $\bar{x} \in R$ ($f'(\bar{x}) = 0$) is a local max point of the problem (P1) if $f''(\bar{x}) < 0$.

IT BOONKRIE NPTEL ONLINE CERTIFICATION COURSE 6

Now to study the optimality conditions of a optimization problem we first start with unconstrained optimization problem, now what is a unconstrained optimization problem let us discuss.

(Refer Slide Time: 16:02)

$$\begin{aligned} & f: \mathbb{R} \rightarrow \mathbb{R} \\ \text{Min } & f(x) \\ \text{s.t. } & x \in \mathbb{R} \end{aligned}$$

Minimizing of $f(x)$ and subject to $x \in \mathbb{R}$ on \mathbb{R} okay we are first taking function from \mathbb{R} to \mathbb{R} if we are optimizing means maximizing or minimize function without any constrained, then we see such problems as uncontained optimization problems. Now to study the optimality condition for such type of problems they start with a function from \mathbb{R} to \mathbb{R} , okay we already know that if we have to minimize a function from \mathbb{R} to \mathbb{R} how we can minimize this function.

(Refer Slide Time: 16:54)

Optimality Conditions

Unconstrained optimization problem
(P1) Min/Max $f(x), x \in R$
(Assume that f is twice differentiable over R).

Necessary Condition: If $\bar{x} \in R$ is a local min or local max point of f over R , then $f'(\bar{x}) = 0$.

Sufficient Condition:

- The point $\bar{x} \in R$ ($f'(\bar{x}) = 0$) is a local min point of the problem (P1) if $f''(\bar{x}) > 0$.
- The point $\bar{x} \in R$ ($f'(\bar{x}) = 0$) is a local max point of the problem (P1) if $f''(\bar{x}) < 0$.

BY ROCKWELL NPTEL ONLINE CERTIFICATION COURSE 6

We first find we first have a necessary condition is if $\bar{x} \in R$ is a local minimum or local maximum point of f over R then always f' of \bar{x} will be 0, we are assuming function is differentiable twice differentiable we assuming a function is twice differentiable, so if \bar{x} is a point of local maximum or local minimum then at that point first derivative is always 0, okay what is a sufficient condition to see weather that point is a point of local maximum, local minimum we find second derivative at that point, at \bar{x} if second derivative at that point is positive then that point is a point of local minimum.

And if secondary derivative at that point is < 0 we call that point has local maximum point, this we already know okay now we come to function from R to R suppose we are taking function from. Function f from R to R .

(Refer Slide Time: 17:54)

$$\begin{aligned} & f: \mathbb{R}^2 \rightarrow \mathbb{R} \\ \text{Min } & f(x) \\ \text{s.t. } & x \in \mathbb{R}^2 \end{aligned}$$

And functions $Rx \in \mathbb{R}^2$ so how to optimize how to find maximum, minimum values of such functions we first find we first have necessary condition what is that.

(Refer Slide Time: 18:10)

Continued...

Now let in the problem (P1), $f : \mathbb{R}^2 \rightarrow \mathbb{R}$.
(i.e. function of two variables).

Necessary Condition: If $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ is a point of local min or local max of f over \mathbb{R}^2 , then

$$\left(\frac{\partial f}{\partial x} \right)_{(\bar{x}, \bar{y})} = \left(\frac{\partial f}{\partial y} \right)_{(\bar{x}, \bar{y})} = 0. \quad (1)$$

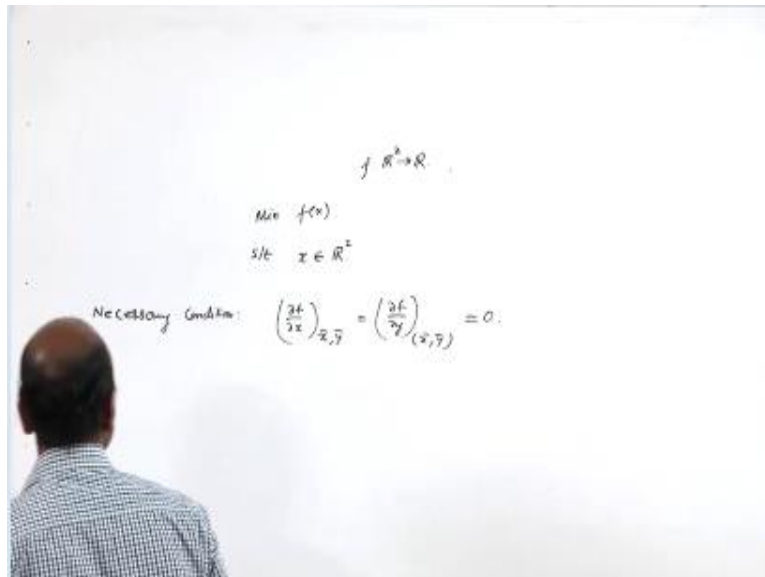
Sufficient Condition:

- If $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ satisfy (1) above and $(\nabla^2 f)_{(\bar{x}, \bar{y})}$ is positive definite then (\bar{x}, \bar{y}) is local min point of f over \mathbb{R}^2 .
- If $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ satisfy (1) above and $(\nabla^2 f)_{(\bar{x}, \bar{y})}$ is negative definite then (\bar{x}, \bar{y}) is local max point of f over \mathbb{R}^2 .

IT BOONEREE NPTEL ONLINE CERTIFICATION COURSE 7

If $\bar{x}, \bar{y} \in \mathbb{R}^2$ is a point of local minimum or local maximum of f over \mathbb{R}^2 then $\partial f / \partial x$ at \bar{x} and $\partial f / \partial y$ at \bar{y} is always 0. The first condition is the necessary condition what is a necessary condition for two variables function.

(Refer Slide Time: 18:31)



Okay that is simply that $\partial f / \partial x$ at \bar{x} \bar{y} = $\partial f / \partial y$ at \bar{x} / \bar{y} must be 0, so there is a first condition and necessary condition now if this condition hold whether that point is a point local minimum or local maximum or straddle point who can we see, so we find Hessian matrix of f at \bar{x} \bar{y} , these are sufficient conditions you see if \bar{x} , \bar{y} belongs \mathbb{R}^2 satisfy this condition.

(Refer Slide Time: 19:11)

Continued...

Now let in the problem (P1), $f : \mathbb{R}^2 \rightarrow \mathbb{R}$.
(i.e. function of two variables).

Necessary Condition: If $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ is a point of local min or local max of f over \mathbb{R}^2 , then

$$\left(\frac{\partial f}{\partial x} \right)_{(\bar{x}, \bar{y})} = \left(\frac{\partial f}{\partial y} \right)_{(\bar{x}, \bar{y})} = 0. \quad (1)$$

Sufficient Condition:

- If $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ satisfy (1) above and $(\nabla^2 f)_{(\bar{x}, \bar{y})}$ is positive definite then (\bar{x}, \bar{y}) is local min point of f over \mathbb{R}^2 .
- If $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ satisfy (1) above and $(\nabla^2 f)_{(\bar{x}, \bar{y})}$ is negative definite then (\bar{x}, \bar{y}) is local max point of f over \mathbb{R}^2 .

IT BOONEREE NPTEL ONLINE CERTIFICATION COURSE 7

That is partial derivative of f respect to $2x$ and partial derivative of x respect to y both are 0 at x bar, y bar and hessian matrix of f at x bar, y bar is positive definite the x bar, y bar is point of local minimum and if it is negative definite then it will be point of local maximum of f/\mathbb{R}^2 okay you see if it is positive definite what does it mean it means function is strictly convex function is convex that means we have a parabolic shape like this okay so certainly we have a point of minima and if we have a if it is negative definite that means function is strictly concave.

A function is concave that means we have this type of parabola and that means that point we are point of local maxima that we can easily visualize okay.

(Refer Slide Time: 20:17)

Continued...

In general, for the problem
Min/Max $f(x)$, $f: R^n \rightarrow R$, the necessary and sufficient condition for local min or local max are as follows:

Necessary Condition: Let $\bar{x} \in R^n$ be a point of local min or local max of f over R^n , then

$$\nabla f(\bar{x}) = 0.$$

Sufficient Condition: Let $\bar{x} \in R^n$ with $\nabla f(\bar{x}) = 0$. If f is a strictly convex (strictly concave) function in a neighbourhood of \bar{x} , then \bar{x} is a local min(local max) point of $f(x)$ over R^n .

Problem Find the points of local max or min(if exists) for the function
 $f(x) = 2 + 2x_1 + 3x_2 - x_1^2 - x_2^2$

NPTEL ONLINE CERTIFICATION COURSE

Now the same conditions we can generalize for a function from R^n to R okay, if we have function R^n to R see.

(Refer Slide Time: 20:27)

$$\begin{aligned} & f: \mathbb{R}^n \rightarrow \mathbb{R} \\ \text{Min } & f(x) \\ \text{s.t. } & x \in \mathbb{R}^n \\ \text{Necessary Condition: } & \nabla f(\bar{x}) = 0 \Rightarrow \left(\frac{\partial f}{\partial x_j} \right)_{\bar{x}} = 0, \quad j=1, 2, \dots, n. \end{aligned}$$

So how can we see it how can you find the necessary and sufficient condition for a point to be local maxima local minima for that we first find the necessary condition will be gradient of f at \bar{x} must be 0 gradient of \bar{x} means this implies $\partial f / \partial x_j$ or i at \bar{x} must be 0 for all j we are having n number of variables okay, and what is your sufficient condition.

(Refer Slide Time: 21:08)

Continued...

In general, for the problem
Min/Max $f(x)$, $f: R^n \rightarrow R$, the necessary and sufficient condition for local min or local max are as follows:

Necessary Condition: Let $\bar{x} \in R^n$ be a point of local min or local max of f over R^n , then

$$\nabla f(\bar{x}) = 0.$$

Sufficient Condition: Let $\bar{x} \in R^n$ with $\nabla f(\bar{x}) = 0$. If f is a strictly convex (strictly concave) function in a neighbourhood of \bar{x} , then \bar{x} is a local min(local max) point of $f(x)$ over R^n .

Problem Find the points of local max or min(if exists) for the function
 $f(x) = 2 + 2x_1 + 3x_2 - x_1^2 - x_2^2$

NPTEL ONLINE CERTIFICATION COURSE

If $\bar{x} \in R^n$ with gradient of $f(\bar{x}) = 0$ if f is strictly convex function in the neighborhood of \bar{x} then \bar{x} is point of local minima and if it is a local minima of f over \bar{x} a strictly convex means hessian matrix is positive definite okay and if hessian matrix is negative definite then \bar{x} will be a point local maxima the same result which we did for a two variable problem the same result we can generalize for n variable problem also okay, now suppose this problem you are having you want to find out a local maxima, local minima for this problem this simple two variable problem we are having, so what is f

(Refer Slide Time: 22:07)

$f = 2 + 2x_1 + 3x_2 - x_1^2 - x_2^2$
 $\frac{\partial f}{\partial x_1} = 0, \quad \frac{\partial f}{\partial x_2} = 0$
 $2 - 2x_1 = 0 \quad ; \quad 3 - 2x_2 = 0$
 $\Rightarrow x_1 = 1 \quad ; \quad x_2 = \frac{3}{2} \quad \text{at } (1, \frac{3}{2})$
 $\nabla^2 f = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$
 $D_1 = -2 < 0$
 $D_2 = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4 > 0$
 $(1, \frac{3}{2}) \rightarrow \text{local max.}$

f area is nothing but $2+2x_1+3x_2 -x_1^2-x_2^2$ okay, where enter should find out its point of local maxima, local minima for this problem. So first we find $\partial f/\partial x_1$ put it equal to 0 and $\partial f/\partial x_2$ and put it equal to 0, this will give us stationery point this condition will give us stationery point, okay. So what is $\partial f/\partial x_1$ it is two $2+2-2x_1=0$ and here where it is $3-2x_2=0$ so this implies $x_1=1$ and this implies $x_2=3/2$ so what is the point, point is $(1, 3/2)$. So $(1, 3/2)$ is the only stationery point for this particular problem.

Now we have to see that whether this point is a point of local maxima or local minima or neither or a settle point that we have to see, so for that we find initial matrix of f at this point, okay here what is initial matrix of f, initial matrix I fill be nothing but it is -2 it is 0,0 and again -2, you find determinants of the leading principle -, D_1 is -2 which is negative you get is this see determinant of 1×1 it is negative, and what is D_2 .

D_2 is determinant of 2×2 which is $-2, 0, 0, -2$ which is again 4 which is greater than 0, so first is first one is negative, second one is positive that means it is alternating in sign that means this is negative definite, this initial matrix negative definite, initial matrix and negative definite means function is strictly concave and that means strictly concave means we have a upper hyperbola

like this and that means that point will be a point of local maxima, so that means the point $(1, 3/2)$ will be a point of local maxima.

So that is how we can easily check whether a point is the point of, whether a point which is the obtained by this stationery condition is the point of local maxima, local minima or neither, okay. Now let us discuss constrained optimization problem first with equality constraints and then we will see with inequality constraints, okay.

(Refer Slide Time: 24:59)

Constrained Optimization Problem with Equality Constraints

Consider the problem:

(P2) Min/Max $f(x)$
subject to: $g_i(x) = 0, i = 1, 2, \dots, m$
where $f: \mathbb{R}^n \rightarrow \mathbb{R}, g_i: \mathbb{R}^n \rightarrow \mathbb{R}$.

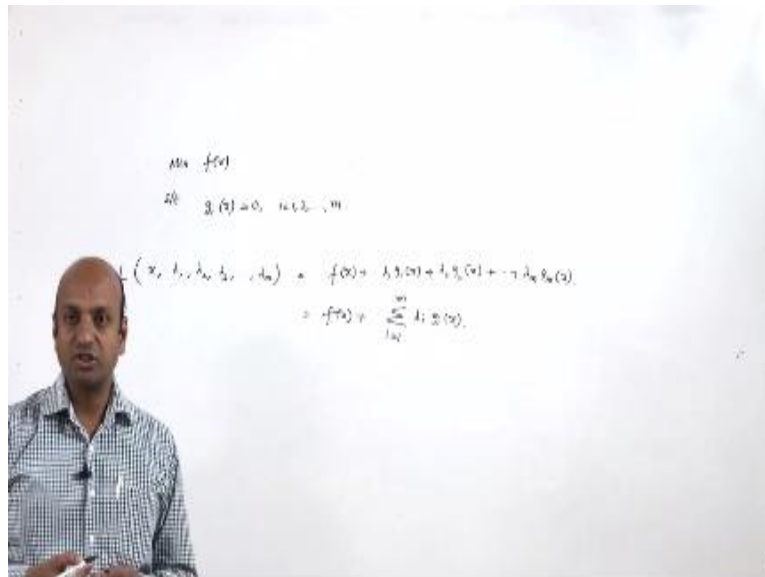
We construct the Lagrange's function or the Lagrangian as:

$$L(x, \lambda_1, \lambda_2, \dots, \lambda_m) = f(x) + \sum_{i=1}^m \lambda_i g_i(x).$$

NPTEL ONLINE CERTIFICATION COURSE

Now what do you mean by constrained optimization problem, constrained optimization problem means optimizing that is maximizing or minimizes function subject to some conditions, some constraints okay. So minimizing or maximizing of $f(x)$ subject to $g_i(x) = 0$ because we are taking only equality constraints here i from 1 to m this is the problem which we are discussing now and how to find out optimal solutions of such problems, optimal solution means the problem we satisfy all the constraints as well as minimize or maximize the objective function, so let us see how to find it let us discuss it.

(Refer Slide Time: 25:52)



So we have the problem suppose they have problem minimizing of $f(x)$ subject to $g_i(x)=0$ i from 1 to m , so this is the constrained optimization problem with equality constraints. Now how to find out the optimal solution for such problems we first define a function L which we call as Lagrange function or Lagrangian, so this Lagrangian will involve all x , all n components of x with λ s and these λ s are called Lagrange multipliers.

But we first define this function which we call as Lagrange function, so what is this function is this is nothing but objective function plus $\lambda_1 g_1 + \lambda_2 g_2$ and so on $\lambda_m g_m(x)$ or it is $f(x) + \sum_{i=1}^m \lambda_i g_i$ so this is the Lagrange function, okay.

(Refer Slide Time: 27:11)



Necessary Condition

Let $\bar{x} \in \mathbb{R}^n$ be a point of local min or local max of the problem (P2), where $m < n$. Let it be possible to choose set of m -variables x_i for which the Jacobian matrix $J = \left[\left(\frac{\partial g_i}{\partial x_j} \right)_{\bar{x}} \right]_{m \times m}$ has an inverse. Then there exists a unique set of Lagrange multipliers $\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_m$ such that

$$\nabla f(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i \nabla g_i(\bar{x}) = 0. \quad (2)$$

That is,

$$\frac{\partial L}{\partial x_j} = 0, j = 1, 2, \dots, n$$
$$\frac{\partial L}{\partial \lambda_i} = 0, i = 1, 2, \dots, m.$$

 01 022023  OPJS ONLINE CERTIFICATION COURSE 10

Now what are the necessary condition, what are necessary condition for this problem let $x_j \in \mathbb{R}^n$ be a point of local minimum or local maximum for the problem P2, P2 is this problem, where $m < n$ okay, let it be possible to choose set of m variables x_i for which Jacobian matrix $(\partial g_i / \partial x_j)_{\bar{x}}$ which is what are $m \times m$ has an inverse, it must be $m \times m$ has an inverse, okay. So this condition is required for in order to prove this I am not discussing the proof here, I am just discussing the conditions, okay.

Then there exists a unique set of Lagrange multiples $\bar{\lambda}_1, \bar{\lambda}_2$ and so on up to $\bar{\lambda}_m$ such that these condition, such that this condition hold. So what in necessary condition, necessary condition is $\nabla f(\bar{x}) + \sum_{i=1}^m$

(Refer Slide Time: 28:17)

The image shows handwritten mathematical work on a whiteboard. At the top right, there is an equation: $\frac{\partial f}{\partial x_j} + \sum_{i=1}^m \lambda_i \frac{\partial g_i}{\partial x_j} = 0, j=1, \dots, n$. Below this, the Lagrangian function is defined as $L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m) = f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x) + \dots + \lambda_m g_m(x)$. This is then simplified to $L = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$. The next line shows the gradient of the Lagrangian with respect to x_j set to zero: $\nabla_x L(x) + \sum_{i=1}^m \lambda_i \nabla_x g_i(x) = 0$. This is followed by a detailed expansion of the gradient components for $j=1$: $\left(\frac{\partial f}{\partial x_1} + \lambda_1 \frac{\partial g_1}{\partial x_1} + \lambda_2 \frac{\partial g_2}{\partial x_1} + \dots + \lambda_m \frac{\partial g_m}{\partial x_1} \right) = 0$. A similar expansion is shown for $j=2$: $\left(\frac{\partial f}{\partial x_2} + \lambda_1 \frac{\partial g_1}{\partial x_2} + \lambda_2 \frac{\partial g_2}{\partial x_2} + \dots + \lambda_m \frac{\partial g_m}{\partial x_2} \right) = 0$.

$\lambda_i \nabla g_i(\bar{x})$ and this $\bar{\lambda}$ must be 0 so this is the necessary condition. Now what just condition means what is gradient $f(\bar{x})$ this is $\partial f/\partial x_1, \partial f/\partial x_2$ and so on $\partial f/\partial x_n$ at $\bar{x} + \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x})$ this is $\partial g_i/\partial x_1, \partial g_i/\partial x_2, \partial g_i/\partial x_m$ this equal to 0, now all component must be 0, okay.

If you take this so that means $\partial f/\partial x_1 + \lambda_1 \partial g_1/\partial x_1 + \lambda_2 \partial g_2/\partial x_1 + \dots + \lambda_m \partial g_m/\partial x_1$ should be 0, the first condition because it takes the first component and you vary i from 1 to m is the first component so you will get this condition equal to 0. Similarly $\partial f/\partial x_2 + \lambda_1 \partial g_1/\partial x_2$ it is all bar okay, and it is all at \bar{x} must be 0, similarly you will get the same condition for x_2 same condition for x_n , so what would it means, this means $\partial f/\partial x_i + \sum_{i=1}^m \lambda_i \partial g_i/\partial x_i$ should be 0 and j varying from 1 to m , okay.

Because the same way whole for $i, j=1$ that is the first condition the same will hold when you put $j=2$ that the second condition from here, okay and similarly this is a same will hold for $j=n$, okay or the same conditions can be obtained from the Lagrange function itself you simply take $\partial L/\partial x_i = 0$ for all i, i from 1 to m and $\partial L/\partial \lambda_j = 0$ for all j .

(Refer Slide Time: 31:17)

Let $f(x)$
Let $g_i(x) = 0, i=1, \dots, m$

$$L(x, \lambda_1, \lambda_2, \dots, \lambda_m) = f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x) + \dots + \lambda_m g_m(x)$$
$$= f(x) + \sum_{i=1}^m \lambda_i g_i(x)$$
$$\frac{\partial L}{\partial x_i} = 0, i=1, 2, \dots, n$$
$$\frac{\partial L}{\partial \lambda_j} = 0, j=1, 2, \dots, m$$

So from these two also we get back to this that condition 1, okay. So these are necessary condition if the point is a point of local minima, okay then this condition 2 must hold, okay. Now what is the sufficient condition, let $\bar{x}, \bar{\lambda}$ which belongs to $\mathbb{R}^n \times \mathbb{R}^m$ exists.

(Refer Slide Time: 31:54)

Sufficient Condition

Let $(\bar{x}, \bar{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^m$ exists such that condition (2) above holds. Suppose $Z(x^*) = \{z \in \mathbb{R}^m : z^T \nabla g(\bar{x}) = 0\}$ and $z^T (\nabla^2 L(\bar{x}, \bar{\lambda})) z > 0$, for all $z \in Z(x^*)$ with $z \neq 0$, then \bar{x} is a local min point of (P2).

Similarly, if $z^T (\nabla^2 L(\bar{x}, \bar{\lambda})) z < 0$, for all $z \in Z(x^*)$ with $z \neq 0$, then \bar{x} is a local max point of (P2).

Questions

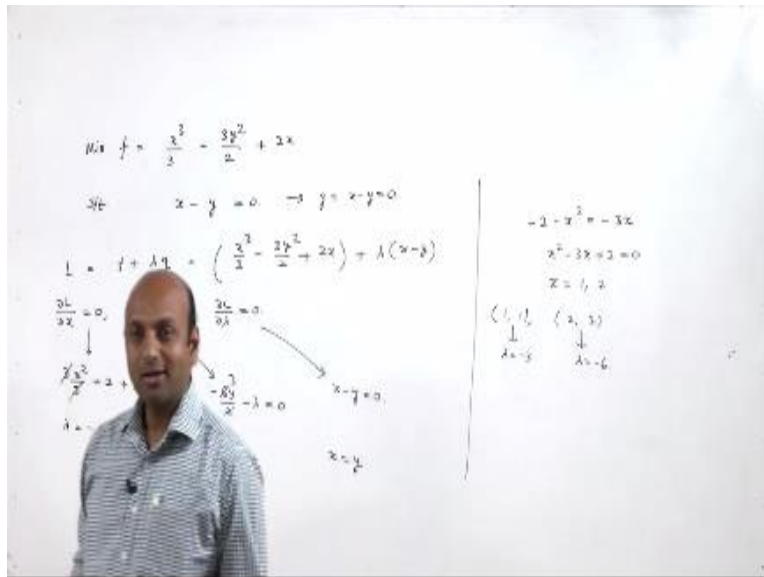
Use the method of Lagrange's multiplier to solve

- Min $\frac{x^3}{3} - \frac{3y^2}{2} + 2x$
subject to: $x - y = 0$
- Optimize $f(x) = 2x_1^2 + x_2^2 + 3x_3^2 + 10x_1 + 8x_2 + 6x_3 - 100$
subject to: $x_1 + x_2 + x_3 = 200$

11

Such that the condition 2 above holds this condition hold. Suppose $Z(x^*)$ consists of all $Z \in \mathbb{R}^m$ such that Z^T gradient of $g\bar{x}=0$ and Z^T of initial matrix of L with $Z>0$ for all $Z \in Z(x^*)$ where Z not equal to 0 then x^* is the point of local, x has the point of local minima of the problem P2, P2 is this problem, okay. So let us discuss this condition with an example, okay this example will clear all this things so let us discuss this fine example.

(Refer Slide Time: 32:47)



Suppose we have the first problem minimizing $f = x^3/3 - 3y^2/2 + 2x$ subject to $x - y = 0$, so one way out is very easy you substitute $x = y$ in this problem it will become a one variable unconstrained optimization problem and you can find out first derivative respect to x put it equal to 0 and then you second derivative to see whether rare point to the point of local maxima or local minima, okay. You simply see you substitute $x = y$, $x = y$ here and $x = y$ here so it will be a one variable unconstrained optimization problem, okay.

You find first derivative put it equal to 0 that will be the stationery point and to check where that point to the point of local maxima or local minima find second derivative at that point if it is greater than 0 means local minima if it is lesser 0 means local maxima. Now if you have solve the same problem using Lagrange functioning how we can do that, so this is simply an illustration so we consider a Langrage function L which is $f + \lambda g$ because only one constraint is involved so it is λg .

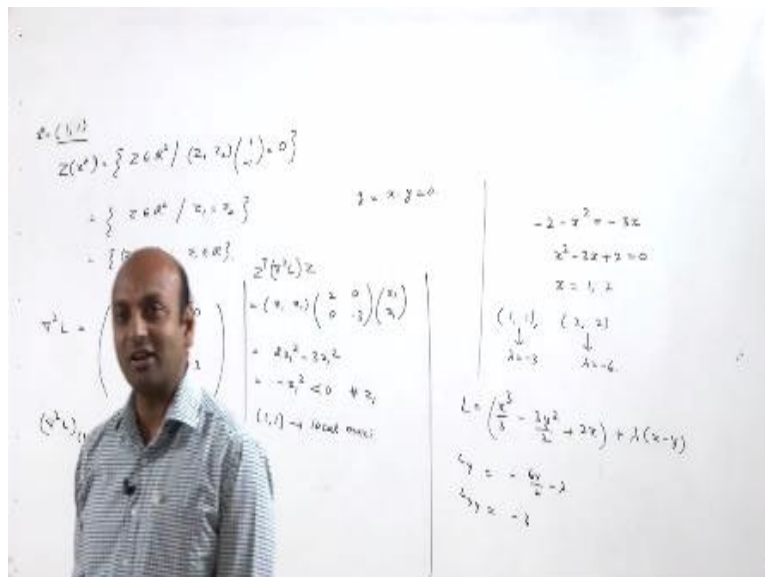
g is this comes front this is $g = x - y = 0$, so this is nothing but $(x^3/3 - 3y^2/2 + 2x) + \lambda(x - y)$, now we find $\partial L / \partial x$ put it equal to 0 and $\partial L / \partial y$ put it equal to 0 and $\partial L / \partial \lambda$ put it equal to 0, by the necessary conditions. What are necessary conditions, we already discussed the necessary condition is $\partial L / \partial x$;

should be 0 for all j and $\partial l / \partial \lambda_i$ should be 0 for all i which are the same condition which is in 2, okay. So what is $\partial l / \partial x$ first we solve this problem this condition it is $3x^2/3 + 2 + \lambda = 0$ what is this condition, this condition is $-6y/2 - \lambda$ should be equal to 0, what is this condition, this condition is $x - y = 0$ okay. So what we obtain from here 3,3 cancel out here 2 and 3 times cancel out so this is nothing but λ will be equals to $-2 - x^2$ from here, from here λ must be equals to $-3y$ and from here $x = y$, okay.

Though you substitute $x = y$ and $\lambda = \lambda$ so what we obtain, we obtain $-2 - x^2 = -3x$ because the $x = y$ and $\lambda = \lambda$ okay so what we obtain it is $x^2 - 3x + 2 = 0$ that means x are 1 and 2. If x is 1 and 2 and $x = y$ that means y is also 1 and 2 so points are (1,1) and (2,2) and when x is 1 or y is 1 λ is -3 so corresponding to this λ is -3 and corresponding to this λ is -6 because you substitute $y = 2$ so $\lambda = -3$, so these are the stationery points, okay.

Now to see whether which point is a point of local minima and which point is a point of local maxima we will apply sufficient condition, so what is sufficient condition let us see. First we find the Z, so what Z indicates.

(Refer Slide Time: 37:22)



First we find for suppose (2,2) or (1,1) okay (1,1) first we see for this point this is x^* for this problem so we will find Z at x^* this is x^* , we will find all those $Z \in \mathbb{R}^2$ here because n is 2 such that Z^T , Z^T means z_1, z_2 and gradient of g_1 what is g here $x-y=0$, constraint is $x-y=0$ what will be its gradient, gradient will be respect to x 1 respect to y $-1=0$, so this implies $z_1=z_2$ okay all $Z \in \mathbb{R}^2$ so what will be $Z(x^*)$ it is of this type Z_1, Z_1 .

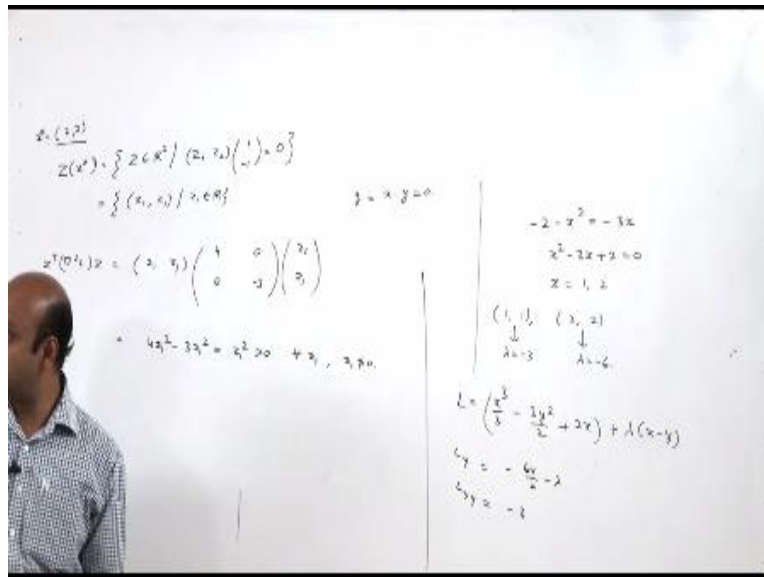
It is z_1, z_2 and $Z_1=Z_2$ that means z_1, z_1 types all the elements will be in Z okay. Now if the first condition hold this condition hold we say that x^* will be point of local minima, if this condition hold we say that this point is a point of local maxima, so let us see we first find initial matrix of L , okay and what is L for this problem we again have to write the L for this problem L will be $x^3/3-3y^2/2$ okay $+2x$ and $+\lambda(x-y)$.

So what is gradient square of L , that will be nothing but we first differential respect to x twice but that is $2x$ okay, now respect to xy which is 0 and with respect to y , it is -6 it is first derivative respect to y is $-6y/2-\lambda$ and when to differentiate this twice respect to y it is nothing but -3 , okay so it is -3 it is second partial derivative respect to y , okay. Now this at (1,1) is nothing but $2,0,0$, -3 okay.

Now you take Z_1^T of this Z_1^T initial matrix of L be Z so it is nothing but let us try to find this value Z^T gradient $\partial^2(L)Z$. Now for this x^* (1,1) Z^* is z_1, z_1 okay z_1, z_1 types and initial matrix is $2,0,0,-3$ again it is z_1, z_1 so it is nothing but $2z_1^2-3z_1^2$ which is $-z_1^2$ which is less than 0 for all z_1 and z_1 not equal to 0, okay.

So this quantity less than 0, this means this points are point of local maxima so that means (1,1) is a point of local maxima, so that is how we can check whether a point by a Lagrange function whether a point is a point of local maxima or local minima. Similarly we will check for (2,2) also when you take (2,2) what will be Z .

(Refer Slide Time: 41:45)



For (2,2) Z will be nothing but this is (2,2) so this would be same, this will remain same and when you find Z^T initial matrix of L with Z so this is nothing but z_1, z_1 . Now initial matrix of L at (2,2) will be it is 4,0,0,-3 and z_1, z_1 because both are same, so it is $4z_1^2 - 3z_1^2$ which is z_1^2 which is positive for all z_1 and z_1 not equal to 0. So this is positive this means the first condition hold that means this is a point of local minima.

So the point (1,1) is a point of local maxima and (2,2) is the point of local minima, so in this way using Lagrange function we can easily check whether point, the point which we are find out is a point of local maxima or local minima. Now similarly we can solve the second problem also using the Lagrange multiplier method, okay so thank you.

For Further Details Contact

Coordinator, Educational Technology Cell

Indian Institute of Technology Roorkee

Roorkee – 247667

E Mail: etcell.iitrke@gmail.com, etcell@iitr.ernet.in

Website: www.iitr.ac.in/centers/ETC, www.nptel.ac.in

Camera

Jithin. K

Graphics

Binoy. V. P

Online & Video Editing

Mohan Raj. S

Production Team

Sarath Koovery

Arun. S

Pankaj Saini

Neetesh Kumar

Jitendar Kumar

Vivek Kumar

Nibedita

Bisoyi

An Educational Technology cell

IIT Roorkee Production

© Copyright All Rights Reserved